Iranian Journal of Chemical Engineering
Vol. 16, No. 2 (Spring 2019), IAChE

Prediction of True Critical Temperature and Pressure of Binary Hydrocarbon Mixtures: A Comparison Between the Artificial Neural Networks and the Support Vector Machine

M. Etebarian, K. Movagharnejad [*]

Faculty of Chemical Engineering, Babol Noshirvani University of Technology, Babol, Iran

ARTICLE INFO

Article history:
Received: 2017-12-08
Accepted: 2019-04-29

Keywords:
Critical Pressure,
Critical Temperature,
Artificial Neural Network,
Support Vector Machine,
Binary Hydrocarbon Mixture,
Particle Swarm Optimization

ABSTRACT

Two main objectives have been considered in this paper: providing a good model to predict the critical temperature and pressure of binary hydrocarbon mixtures, and comparing the efficiency of the artificial neural network algorithms and the support vector regression as two commonly used soft computing methods. In order to make a fair comparison and to achieve the highest efficiency, a comprehensive search method is used in neural network modeling and a particle swarm optimization algorithm is applied to SVM modeling. To compare the accuracy of the models, various criteria such as ARD, MAE, MSE, RAE, and R^2 are used. The simulation results show that the ARD for the prediction of the true critical temperature and pressure of the binary hydrocarbon mixtures for the final optimized ANN-based model are equal to 0.0161 and 0.0387, respectively. The corresponding ARD values for the SVM-based model are equal to 0.0086 and 0.0091 for critical temperature and pressure, respectively. Simulation results show that although both models have very high predictive accuracy, the SVM has higher learning speed and accuracy than ANN.

1. Introduction

Critical properties are considered as some of the most important thermodynamic properties of pure chemical compounds and mixtures with various applications to different industrial fields [1]. Knowing the critical properties is necessary essential for estimating other thermodynamic properties or using the common equations of state [2]. Critical properties are defined as characteristics such as temperature and pressure in certain conditions, where there is no distinction between the liquid and vapor phase, and this point is known as the critical state or critical point. Since the measurement of these properties is not always possible, several methods have been established to predict the critical properties of pure components and their mixtures [3]. Traditional methods for the estimation of pure component critical constants usually consist of group, bond or atomic contribution methods [4]. These methods consider the molecules as an arrangement of neighboring groups, bonds or atoms. Recently, other alternative methods such as quantitative

[*]Corresponding author: movagharnejad@yahoo.com

structure-property relationships (QSPR) have been also developed. These methods estimate the critical properties in relation to other structural properties such as molecular weight or normal boiling point. The main advantage of the QSPR methods is that there is no need to know the exact molecular structure of the pure compounds.

In more recent years, some other researchers have applied soft computing methods to estimate the critical constants of pure compounds. Most of these researchers have claimed that the soft computing mehods are more accurate than the traditional estimation group contribution or QSPR methods; however, since these methods usually need a minimum of similar experimental data, their claims have to be considered with caution and need to be approved by further independent investigations.

As most engineering applications deal with mixtures, the critical properties of mixtures are also investigated by several research studies. In addition to experimental methods, other methods such as equations of state and molecular simulation models are also used to estimate the critical properties of mixtures. These computational methods usually require a full understanding of the mixture, which is not allways possible [3]. Due to high costs, laboratory testings are not always possible and as the critical point of mixture deponds on its composition, the experimental procedure may be extra expensive and time-consuming for mixtures.

Although many researchers have worked on this subject, no general and reliable solution has been published, and further researches are still needed. For example, Najafi et al. [5] presented a theoretical method for predicting the critical temperature of multi-component mixtures, in which they tried to improve the Chueh-Prausnitz theory. The absolute relative deviation (ARD) was equal to 2.058 % and 1.800 % for critical temperature and pressure of selected binary mixtures, respectively. The error for their model was relatively high and did not seem to be satisfactory for most engineering applications. On the other hand, although soft computational methods require specific constraints and, sometimes, high computational time, they can be efficiently and accurately presented due to the use of efficient learning methods and have been widely applied to different cases. The accuracy of computational soft computing models depends on the availability of experimental data and the efficiency of the modeling tools [6,7]. Najafi et al. [8] also categorized the predictive methods for critical properties into two main categories of fast estimation and rigorous methods, and examined the most common methods of each of category. They have also reported that the rigorous methods are usually preferred to the fast estimation methods because of the generability and high accuracy; however, they have also listed some of their disadvantages such as method complexity and the need for mixing rules and adjustable parameters. These limitations have limited the use of these rigorous methods in real industrial processes.

Fast estimation methods are in fact empirical equations, which are fitted by experimental true critical data. Their accuracy usually depends on two main factors of experimental data availability and the suitability of the original mathematical formula used to fit the experimental data. The original mathematical formulas such as Li or Chueh-Prausnitz method are simple fitted equations in terms of component critical properties and mole fractions. Since no

simple linear relation suitably fits the experimental data, these mathematical formulas experienced a continuous process of complexity without any fair accuracy improvement. The next generation of fast estimation methods tried to improve the accuracy by means of group contribution technique, which was a very common technique for prediction of critical constants of pure components; a detailed report of them may be found elsewhere [8].

Since there is usually no linear relationship between the critical properties of the mixtures and the mole fractions of their pure components, the selected computational models must have the ability to map the inputs and outputs with non-linear relationships [8].

Traditional computational methods such as regression analysis can be used to predict the true critical properties of mixtures; however, nowadays, intelligent computational methods have been proposed that can provide better mapping accuracy with higher speed. For example, Zhao et al. [9] used the Multiple Linear Regression (MLR) and Support Vector Machine (SVM) to predict the liquid iodine viscosity, and reported that MLR and SVM showed 24.2 % and 3.95 % average absolute relative deviation (AARD) errors, respectively. Therefore, their work indicated that the SVM showed much higher accuracy than the MLR. Another example is the work of Mehdizadeh and Movagharnejad [10], which compared the accuracy of seven semi-empirical equations and artificial neural networks in predicting the solubility of various components in supercritical carbon dioxide. In this paper, they showed that their proposed ANN-based model was much more accurate than the best empirical correlations. Today, various intelligent methods such as artificial neural networks, fuzzy logic, support vector machine, support vector regression, and so on are used to generate predictive models.

Since the support vector machine is a smart computational method, its use in chemical engineering research has increased in recent years. Hayer et al. [11] used SVM and cubic plus association (CPA) equations of state to predict the density of normal alkanols, and showed that the SVM based model and the CPA thermodynamic model had 0.06 % and 1.36 % AARD errors, respectively. Their work suggested that SVM could predict the thermodynamic properties more accurately than CPA equation of state. Although the SVM has a high predictive power, it should solve a large number of second-order equations and is very time consuming; thus, in this paper, the least squared SVM (LSSVM) is used to reduce the computational time and increase the execution speed.

Among various prediction tools, Lashkarbolooki et al. used artificial neural networks as a powerful tool that can detect complex interactions between inputs and outputs to develop the prediction device for two-component heat capacity of ionic liquids mixtures. Their work consisted of multilayer supervised perceptron networks, because this type of network is suitable for determining the non-linear relationships between the inputs and the outputs using previous training data [12]. The optimal number of neurons, the number of hidden layers, and the transfer function were specified during the training process.

Sabzevari and Moosavi argued that, due to the highly nonlinear relationship between the physical and thermodynamic properties, artificial neural networks may provide an appropriate algorithm for learning the

relationship between the input and the output vectors and estimating each thermodynamic property such as density [13]. They introduced this method as a suitable alternative way for predicting the alkali metal density.

In 2013, Hosseini-Nasab et al. [6] used adaptive network-based fuzzy inference system (ANFIS) to predict the vapor liquid equilibrium (VLE) of binary systems. Fuzzy logic reduces the complexity of data modeling and enables the artificial neural networks to capture the behavior of more complex systems. Therefore, ANFIS is developed using the capabilities of these two computational methods. Using experimental data, they have shown that the prediction based on the ANFIS model is more accurate than the conventional Redlich-Kwang–Soave (RKS) equation of state.

Ansari and Gholami [7] tried to introduce an intelligent model for the crude oil saturation pressure. They tried to determine the relationship between saturation pressure and the composition of hydrocarbon and non-hydrocarbon components of the crude oil using a support vector regression and optimization techniques. They also used various optimization techniques such as genetic algorithm, colonial competition algorithm, and particle swarm optimization algorithm to optimize the model parameters for the SVM model.

Since ANN is the most widely used smart computing device and the SVM is one of the most successful computing techniques in recent years, it was decided to use these two smart computational methods in this study. On the other hand, each molecule has been divided into a number of standard functional groups. These numbers are considered as model inputs; thus, our models are really a combination of intelligent and group contribution methods. Two different models have been developed and optimized to predict the critical pressure and temperature of binary hydrocarbon mixtures. Finally, the results of these two models have been analyzed and compared with each other. Experimental data have been gathered and used to train these two smart soft computing models, and the particle swarm optimization (PSO) algorithm was used to optimize the support vector regression parameters. Other separated experimental data were used for testing and optimizating the trained soft computing models.

2. Theory

2.1. Artificial neural networks

Artificial Neural Networks are considered as a simple model of human brain's neuronal function simulating the learning ability. In machine learning problems, ANN can avoid the regression weaknesses by identifying non-linear and complex relationships between the input and output data [14]. Each artificial neural network has three main layers called input, hidden, and output. The hidden layer may consist of several sub-layers. The non-linear relationships between these layers are constructed through the parallel structures of the neurons determining the performance of the grid, the weight of each neuron, and the number of the hidden layers [14].

Artificial neural networks may be used in different learning problems with observer, without observer, and for reinforcement. Observer learning is usually used in prediction problems that take advantage of past experiences and knowledge. In these problems, the model can be shown with a set of Input Dataset to find a function capable of mapping the input data to the corresponding

tags. The cost function is used to determine the mapping quality, and the mapping errors are detected on the basis of the previous knowledge. Multilayer Perceptron Neural Network (MLP) is one of the most common neural architectures used to predict the different thermodynamic properties. One or more hidden layers are considered for training this kind of ANN, and each node in a layer connects to all of the nodes of the next layer. Then, a random weight is assigned to each neuron, and the array of weights with the lowest cost is considered as the optimal model. Finally, the model with the lowest cost is selected as the trained model. Since the MLP is known to be a feed–forward network, the data moves in one direction in the training section and this fact justifies the high speed of MLP during the testing phase. More information may be available on the artificial neural network modeling and its different architectures elsewhere [14].

2.2. Support vector machine

Support vector regression is known to be one of the most powerful methods of learning with observer that uses statistical learning theory and is successfully applied to classification, regression, and data analysis problems. The support vector regression is considered as a special kind of support vector machine. A SVM model is developed on the basis of the structural risk minimization criterion using an instruction set.

In the SVM, at first, the input data are mapped to a higher attribute space so that a decision and linear relationship can be found in this space between the data. Solving complex second-order equations by this method requires a lot of time. Thus, the least square support vector regression (LSSVR) algorithm is presented to simplify the complex problems and reduce them to first-order equations. Further details may be found elsewhere [15]. In order to achieve the maximum efficiency in a LSSVR model, you have to set the model parameters with optimal values.

2.3. Particle swarm optimization algorithm

By setting the model parameters, the model output may be fitted to the experimental data as far as possible [16]. Thus, achieving the proximity to the experimental values should be considered according to each specific case study. Comprehensive search can be used for any case of limited range and number of adjustable parameters; otherwise, modern heuristic optimization techniques have to be used. Various optimization algorithms such as evolutionary computation, simulated annealing, tabu search, genetic algorithm (GA), and particle swarm optimization (PSO) have been applied to different applications, recently. PSO has numerous advantages such as the accuracy and speed [17]. Thus, we have used this technique to optimize the SVM. Algorithms such as GA are of discrete nature and their accuracy ceases for continuous applications; however, PSO has a continuous structure and is more suitable for predicting the continuous values such as critical temperature and pressure of mixtures [18].

The PSO is a simplified model of the process of finding food by a bundle of birds, fish, or animals as an heuristic optimization approach [17]. The overall strategy of the PSO is that the model parameters may be considered as simple components randomly defined in the problem space to calculate the accuracy of the model [19]. Each component, which is called a particle, is then displaced by one or more members of the Swarm community. The movement history and the

best place location are memorized. All particles are displaced in each subsequent iteration, and this process continues until all particles are concentrated in a limited location. A single particle cannot solve the problem; however, the interaction between each particle and other particles makes the optimization process meaningful [20]. The PSO mathematics and a full description of its types are given elsewhere [19, 20].

3. Methods and experimental data

Due to the importance of the critical point in the industrial applications, especially the hydrocarbon processing, the precise determination of the true critical point of hydrocarbon mixtures is very crucial. In this paper, an attempt has been made to use two intelligent computing methods to create two efficient and accurate models for predicting the critical temperature and pressure of binary hydrocarbon mixtures. ANN and SVM are used to generate models due to their accuracy and reliability, which have been also reported by several studies in recent years. Of course, there are also many different empirical and semi-empirical correlations published in the scientific literature; however, our goal is to compare two more common modern intelligent techniques with each other. The idea behind this research is that the modern soft computing intelligent techniques have to be more deeply investigated before being widely used in industrial applications.

3.1. Dataset collection

The accuracy of the collected training and test data and the comprehensiveness of the database have a great impact on the accuracy and performance of the intelligent models. The experimental critical data (critical temperature and pressure) were collected from different published papers. The collected database contains the first component mole fraction, the critical temperature of the mixed components, the molecular weight of the mixed components, acentric factor of the mixed components, the number of carbon atoms of the mixed components, the normal boiling point of the mixed components, and the family type of the components. Physiochemical and critical properties of the pure components used in this study are presented in Table 1. The hydrocarbons of Table 1 consist of several different families including alkanes, cycloalkanes, alkenes, and so on. In this respect, 163 binary mixtures with 1646 data points were collected for the prediction of critical temperature, and 123 binary mixtures with 1090 data points were collected for the prediction of critical pressure. The mole fraction range, critical temperature and pressure range, and the references of the selected mixtures for predicting critical temperature and pressure of the binary hydrocarbon mixtures are presented in Tables 2 and 3, respectively.

Table 1
Physiochemical and critical properties of pure components [20].

Methane	190.56	4.599	111.661	0.0115	16.04
Ethane	305.322	4.872	184.55	0.0995	30.07
Propane	369.827	4.248	231.111	0.1523	44.10
Butane	425.122	3.796	272.65	0.2002	58.12
Pentane	469.7	3.37	309.222	0.2515	72.15
Hexane	507.6	3.025	341.877	0.3013	86.18
Heptane	540.2	2.74	371.577	0.3495	100.20
Octane	568.7	2.49	398.827	0.3996	114.23

Table 1
Physiochemical and critical properties of pure components [20].

Nonane	594.6	2.29	423.972	0.4435	128.26
Decane	617.7	2.109	447.305	0.4923	142.29
Undecane	639	1.946	469.077	0.5303	156.31
Dodecane	658	1.82	489.472	0.5764	170.34
Tridecane	675	1.68	508.616	0.6174	184.37
Tetradecane	693	1.57	526.727	0.6430	198.39
Pentadecane	708	1.48	543.833	0.6863	212.42
Hexadecane	723	1.399	560.016	0.7174	226.45
Cyclopentane	511.761	4.502	322.4	0.1959	70.13
Cyclohexane	553.577	4.073	353.872	0.2096	84.16
Cycloheptane	604.3	3.84	391.938	0.243	98.19
2-Methylpropane	408.138	3.648	261.427	0.1808	58.12
2-Methylbutane	460.138	3.381	300.994	0.2275	72.15
Neopentane	433.777	3.199	282.65	0.1964	72.15
Methylcyclohexane	572.188	3.471	374.083	0.2350	98.19
Ethylene	282.338	5.04	169.411	0.0865	28.05
Propene	365.572	4.665	225.461	0.1398	42.08
Acetylene	308.322	6.139	189.35	0.1873	26.04
Benzene	562.161	4.898	353.238	0.21	78.11
2,2-Dimethylbutane	488.888	3.081	322.877	0.2350	86.18
2,3-Dimethylbutane	499.977	3.127	331.127	0.2461	86.18
2-Methylpentane	497.5	3.006	333.411	0.2774	86.18
3-Methylpentane	504.427	3.124	336.422	0.2737	86.18
Toluene	591.8	4.106	383.777	0.2621	92.14
Ethylbenzene	617.2	3.606	409.35	0.3026	106.17
Isopropylbenzene	631.1	3.209	425.561	0.3258	120.19
Methylcyclopentane	532.788	3.785	344.961	0.2302	84.16
Cis-Decalin	702.25	3.24	468.966	0.2939	138.25
2-Methylheptane	568.7	2.484	390.8	0.3772	114.23
Cyclooctane	647.2	3.57	423.838	0.2904	112.22
O-Xylene	630.327	3.734	417.577	0.3104	106.17
2,2-Dimethylpentane	520.5	2.773	352.338	0.2879	100.20
2,2-Dimethylhexane	549.8	2.53	379.988	0.3378	114.23
2,2-Dimethylheptane	576.8	2.35	405.838	0.3899	128.26
Butane	419.95	4.043	266.9	0.1905	56.11
2-Methyloctane	586.75	2.29	416.427	0.4212	128.26
2-Methylhexane	530.372	2.734	363.2	0.3277	100.20
Naphthalene	748.35	4.051	491.144	0.3022	128.17
Phenanthrene	869.25	2.87	610.027	0.4695	178.23
Anthracene	873	2.87	615.177	0.4857	178.23
2,2,4-Trimethylpentane	543.961	2.568	372.388	0.3022	114.23

Table 2
Critical temperature, mole fraction, and references of selected binary mixtures.

NO.	Component (1)	Component (2)	Range of X	Range of T_C (K)	Ref.
1	Methane	Ethane	0-1	305.32-190.56	[21]
2	Methane	Propane	0-0.9926	369.83-192.3	[21]
3	Methane	Butane	0-1	425.12-190.56	[21]
4	Methane	Pentane	0-1	469.7-190.56	[21,22]
5	Methane	Hexane	0.625-0.90	423.15-263	[21]
6	Methane	Heptane	0-1	540.2-190.56	[21]
7	Methane	Octane	0.787-0.798	323.15-298.15	[21]

Table 2
Critical temperature, mole fraction, and references of selected binary mixtures.

NO.	Component (1)	Component (2)	Range of X	Range of T_C (K)	Ref.
8	Methane	Nonane	0.885-0.86	348.15-273.15	[21]
9	Methane	Decane	0-1	617.7-190.56	[21]
10	Ethane	Propane	0-1	369.83-305.41	[21,23]
11	Ethane	Butane	0-1	425.12-305.32	[21,23]
12	Ethane	Pentane	0-1	469.7-305.32	[21,23]
13	Ethane	Hexane	0-1	507.6-305.32	[23]
14	Ethane	Heptane	0-1	540.2-305.32	[21,23]
15	Ethane	Octane	0-1	568.7-305.32	[23]
16	Ethane	Nonane	0-1	594.6-305.3	[23]
17	Ethane	Decane	0-0.9958	617.7-309.4	[21,23]
18	Propane	Butane	0-1	425.15-369.83	[21,23]
19	Propane	Pentane	0-1	469.7-369.83	[23]
20	Propane	Hexane	0-1	507.6-369.83	[23]
21	Propane	Heptane	0-1	540.2-369.83	[23]
22	Propane	Octane	0-1	568.7-369.83	[23]
23	Propane	Decane	0-1	617.7-369.83	[23]
24	Butane	Pentane	0-1	469.7-425.12	[21,23]
25	Butane	Hexane	0-1	507.6-425.12	[21,23]
26	Butane	Heptane	0-1	540.2-425.12	[21]
27	Butane	Octane	0-1	568.7-425.12	[21,23]
28	Butane	Decane	0-1	617.7-425.12	[23]
29	Pentane	Hexane	0-1	507.6-469.7	[23]
30	Pentane	Heptane	0.255-0.747	526.2-492.2	[21]
31	Pentane	Octane	0-1	568.7-469.7	[23]
32	Pentane	Nonane	0-1	594.6-469.7	[21,23]
33	Pentane	Decane	0-1	617.7-469.7	[23]
34	Pentane	Undecane	0.212-0.895	619.7-506.2	[21]
35	Pentane	Dodecane	0.201-0.898	639.2-507.4	[21]
36	Pentane	Tridecane	0.287-0.914	645.2-518.2	[21]
37	Hexane	Heptane	0.1-0.9	537.5-511.4	[21]
38	Hexane	Octane	0.1-0.9	564.5-515.8	[21]
39	Hexane	Decane	0.1-0.9	609.6-524.4	[21]
40	Hexane	Tridecane	0.1-0.9	666.2-542.2	[21]
41	Hexane	Tetradecane	0.1-0.9	689.2-548.2	[21]
42	Heptane	Octane	0.1-0.9	566.4-543.5	[21]
43	Nonane	Tridecane	0.1-0.9	669.6-605.5	[21]
44	Decane	Dodecane	0.1-0.9	655.3-621.9	[21]
45	Decane	Pentadecane	0.03-0.06	705.7-704.9	[24]
46	Methane	Cyclohexane	0.608-0.765	444.3-294.3	[21]
47	Ethane	Cyclohexane	0-1	553.8-305.32	[23]
48	Pentane	Cyclohexane	0-1	553.8-469.7	[23]
49	Pentane	Cyclopentane	0.1-0.9	507.64-474.3	[21]
50	Hexane	Cyclopentane	0.1-0.9	511.6-508.38	[21]
51	Hexane	Cyclohexane	0.1-0.9	549.12-512.37	[21]
52	Heptane	Cyclopentane	0.1-0.9	538.75-516.53	[21]
53	Heptane	Cyclohexane	0-1	554.1-540	[21,25]
54	Octane	Cyclopentane	0.1-0.9	565.92-521.88	[21]
55	Octane	Cyclohexane	0.1-0.9	568.34-556.48	[21]
56	Nonane	Cyclopentane	0.1-0.9	590.25-527.63	[21]
57	Nonane	Cyclohexane	0.1-0.9	592.04-560.86	[21]
58	Decane	Cyclohexane	0.1-0.9	614.3-565.1	[21]
59	Tridecane	Cyclohexane	0.1-0.9	670.6-581.6	[21]
60	Methane	2-Methylpropane	0.337-0.69	377.6-310.9	[21]
61	Methane	2-Methylbutane	0.15-0.3339	449.2-429.2	[21]
62	Methane	Neopentane	0-0.845	298.15-211.32	[21]
63	Methane	Methylcyclohexane	0.91-0.93	255.4-199.8	[21]
64	Methane	Ethylene	0.11-0.85	273.2-195.2	[21]

Table 2
Critical temperature, mole fraction, and references of selected binary mixtures.

NO.	Component (1)	Component (2)	Range of X	Range of T_C (K)	Ref.
65	Ethane	Ethylene	0.285-0.44	292.5-290	[21]
66	Ethane	Propene	0.0525-0.9502	362.7-310.9	[21]
67	Ethane	Acetylene	0.27-0.68	297-293.11	[21]
68	Ethane	Benzene	0-1	562-305.3	[21,26]
69	Ethane	2-Methylpropane	0-1	407.8-305.32	[23]
70	Propane	2-Methylbutane	0.101-0.899	454.3-382.5	[21]
71	Propane	Acetylene	0.166-0.86	364.8-313.7	[21]
72	Propane	Benzene	0-1	562.15-364.85	[26,27]
73	Propane	2,2-Dimethylbutane	0-1	489-369.83	[23]
74	Propane	2,3-Dimethylbutane	0-1	500-369.83	[23]
75	Propane	Neopentane	0-1	433.8-369.83	[23]
76	Propane	Cumene	0-1	631.35-364.85	[27]
77	Propane	2-Methylpentane	0-1	497.7-369.83	[23]
78	Propane	3-Methylpentane	0-1	504.6-369.83	[23]
79	Pentane	Neopentane	0.175-0.791	462.7-440.6	[21]
80	Pentane	Benzene	0-1	562-469.8	[26]
81	Pentane	Toluene	0-1	591.8-469.8	[21,26]
82	Pentane	Ethylbenzene	0-1	617.2-469.8	[26]
83	Pentane	Isopropylbenzene	0-1	631-469.8	[26]
84	Pentane	2,2-Dimethylpropane	0-1	469.7-433.8	[23]
85	Hexane	Neopentane	0.154-0.721	490.8-446.9	[21]
86	Hexane	Methylcyclopentane	0.1-0.9	530.5-511	[21]
87	Hexane	2,2-Dimethybutane	0.1009-0.8569	504.7-491	[21]
88	Hexane	2,3-Dimethybutane	0.3266-0.7872	505.1-501.7	[21]
89	Hexane	3-Methylpentane	0.1569-0.803	506.8-504.5	[21]
90	Hexane	Cis-Decalin	0.1-0.9	691.7-532	[21]
91	Hexane	Acetylene	0.335-0.625	473-413	[21]
92	Hexane	Benzene	0.1-0.9	554.5-511	[21]
93	Hexane	Toluene	0-1	591.8-507.7	[21,26]
94	Hexane	Ethylbenzene	0-1	617.2-507.7	[21,26]
95	Hexane	Isopropylbenzene	0-1	631-507.7	[26]
96	Hexane	2,2-Dimethylpropane	0-1	507.6-433.8	[23]
97	Hexane	Hexadecane	0-1	723-507.6	[23]
98	Heptane	Ethylene	0.0301-0.9296	535.4-301.5	[21]
99	Heptane	Benzene	0-1	562-540.2	[21,26]
100	Heptane	Ethylbenzene	0-1	617.2-540.2	[26]
101	Octane	2-Methylheptane	0.1978-0.7214	565.8-561.8	[21]
102	Octane	Benzene	0-1	568.7-562	[21,26]
103	Octane	Ethylbenzene	0-1	617.2-568.7	[26]
104	Nonane	Benzene	0-1	594.6-562	[26]
105	Decane	Benzene	0-1	617.7-562	[21,26]
106	Tridecane	Benzene	0-1	675-562	[26]
107	Hexadecane	Benzene	0-1	723-562	[26]
108	Undecane	Benzene	0-1	639-562	[26]
109	Cumene	Benzene	0-1	631.35-562.15	[27]
110	Cycloheptane	Benzene	0-1	604.2-562	[26]
111	Cyclooctane	Benzene	0-1	647.2-562	[26]
112	2,2-Dimethylpropane	Cyclohexane	0-1	553.8-433.8	[23]
113	Cyclopentane	Cyclohexane	0.1-0.9	550.5-517.4	[21]
114	Cyclopentane	Methylcyclopentane	0.1-0.9	531.4-514.7	[21]
115	Cyclopentane	Methylcyclohexane	0.1-0.9	568-520.1	[21]
116	Cyclopentane	Cycloheptane	0.2357-0.7648	583.9-553.4	[21]
117	Cyclopentane	Cyclooctane	0.283-0.8626	621.5-536.1	[21]
118	Cyclohexane	Neopentane	0.254-0.824	533.2-464.8	[21]
119	Cyclohexane	Methylcyclopentane	0.1-0.9	552.1-535.5	[21]
120	Cyclohexane	Methylcyclohexane	0.1-0.9	570.8-556.3	[21]
121	Cyclohexane	Cycloheptane	0.2733-0.7517	589.7-565.7	[21]

Table 2
Critical temperature, mole fraction, and references of selected binary mixtures.

NO.	Component (1)	Component (2)	Range of X	Range of T_C (K)	Ref.
122	Cyclohexane	Cyclooctane	0.1632-0.8677	635.7-566.3	[21]
123	Cyclohexane	Cis-Decalin	0.1-0.9	695.2-575.9	[21]
124	Cycloheptane	Cyclooctane	0.201-0.7603	637.2-613.1	[21]
125	Methylcyclopentane	Methylcyclohexane	0.1-0.9	569-537.8	[21]
126	Cis-Decalin	Benzene	0.1-0.9	695.8-582.1	[21]
127	Cis-Decalin	Ethylbenzene	0.1-0.9	697.1-626.4	[21]
128	Cis-Decalin	O-Xylene	0.1-0.9	698-637.4	[21]
129	2,2-Dimethylbutane	2,2-Dimethylpentane	0.1-0.9	517.99-492.48	[21]
130	2,2-Dimethylbutane	2,2-Dimethylhexane	0.1-0.9	545.5-497	[21]
131	2,2-Dimethylbutane	2,2-Dimethylheptane	0.1-0.9	571.07-502.43	[21]
132	2,2-Dimethylpentane	2,2-Dimethylhexane	0.1-0.9	547.43-524.02	[21]
133	2,2-Dimethylpentane	2,2-Dimethylheptane	0.1-0.9	572.07-527.91	[21]
134	2,2-Dimethylhexane	2,2-Dimethylheptane	0.1-0.9	574.29-553.18	[21]
135	2-Methylpentane	2-Methylhexane	0-1	530.4-497.7	[21,23]
136	2-Methylpentane	2-Methylheptane	0-1	559.7-497.7	[21,23]
137	2-Methylpentane	2-Methyloctane	0-1	582.87-497.75	[21,23]
138	2-Methylhexane	2-Methylheptane	0-1	559.6-530.43	[21,23]
139	2-Methylhexane	2-Methyloctane	0-1	582.87-530.43	[21,23]
140	2-Methylheptane	2-Methyloctane	0-1	582.87-559.7	[21,23]
141	Ethylene	Propene	0.117-0.914	358-293	[21]
142	Ethylene	Acetylene	0.54-0.82	288.7-278.7	[21]
143	Propene	Butene	0.217-0.824	407.6-377.6	[21]
144	Propene	Acetylene	0.114-0.885	358.1-313.7	[21]
145	Neopentane	Benzene	0-1	562-433.8	[27]
146	Benzene	Cyclopentane	0.1-0.9	556.9-516.4	[21]
147	Benzene	Cyclohexane	0.1-0.9	558.21-552.6	[21]
148	Methylcyclopentane	Benzene	0-1	562-532.7	[26]
149	Methylcyclohexane	Benzene	0-1	572.1-561.9	[26]
150	Ethylene	Benzene	0-1	563.65-282.75	[21,28]
151	Toluene	Benzene	0.1-0.98	593.7-564.7	[21]
152	Benzene	Ethylbenzene	0.1-0.9	613.5-570.6	[21]
153	Benzene	O-Xylene	0.1-0.9	625.3-572.4	[21]
154	Benzene	Naphthalene	0.318-0.901	705.1-586.4	[21]
155	Toluene	Neopentane	0.331-0.321	481.7-481.65	[21,29]
156	Cyclohexane	Toluene	0-1	591.8-553.8	[26]
157	Toluene	Ethylbenzene	0.1-0.9	615.5-596	[21]
158	Toluene	O-Xylene	0.1-0.9	627-597	[21]
159	Ethylbenzene	O-Xylene	0.1-0.9	628.7-618.8	[21]
160	Naphthalene	Cyclohexane	0.215-0.54	671.7-604.2	[21]
161	Naphthalene	Ethylene	0.17-0.2886	540.2-325.25	[21]
162	Benzene	Phenanthrene	0.709-0.894	711.7-621.2	[21]
163	Benzene	Anthracene	0.897-0.952	621.7-589.7	[21]

Table 3
Critical pressure, mole fraction, and reference of selected binary mixtures.

NO.	Component (1)	Component (2)	Range of x_1	Range of P_C (MPa)	Ref.
1	Methane	Ethane	0.05-0.99125	8.31-4.722	[21]
2	Methane	Propane	0.1-0.9926	10.16-4.747	[21]
3	Methane	Butane	0.3602-0.9214	13.26-7.17	[21]
4	Methane	Pentane	0.295-0.9804	16.93-5.83	[21,22]
5	Methane	Hexane	0.625-0.9	20.8-12.16	[21]
6	Methane	Heptane	0.441-0.894	24.88-13.14	[21]
7	Methane	Octane	0.787-0.798	27.34-26.65	[21]
8	Methane	Nonane	0.86-0.885	32.3-31.7	[21]
9	Methane	Decane	0.739-0.905	37-20.06	[21]
10	Ethane	Propane	0-1	5.19-4.25	[21,30]

Table 3
Critical pressure, mole fraction, and reference of selected binary mixtures.

NO.	Component (1)	Component (2)	Range of x_1	Range of P_C (MPa)	Ref.
11	Ethane	Butane	0-1	5.813-3.79	[21,30]
12	Ethane	Pentane	0.481-0.9952	6.78-4.99	[21,31]
13	Ethane	Hexane	0.525-0.9935	7.9-5.12	[21,31]
14	Ethane	Heptane	0.2654-0.9913	8.871-4.7	[21,31]
15	Ethane	Octane	0.9748-0.9959	6.42-5.15	[31]
16	Ethane	Nonane	0.9784-0.9928	6.57-5.47	[31]
17	Ethane	Decane	0.698-0.9958	11.82-5.31	[21]
18	Propane	Butane	0.12-0.926	4.36-3.91	[21,32]
19	Propane	Pentane	0.147-0.819	4.58-3.713	[21]
20	Propane	Hexane	0.1435-0.9218	4.99-3.706	[21]
21	Propane	Heptane	0.214-0.909	5.388-3.606	[21]
22	Propane	Octane	0.214-0.959	5.96-3.544	[21]
23	Propane	Decane	0.712-0.987	7.08-4.67	[21]
24	Butane	Pentane	0.1-0.9	3.784-3.44	[21]
25	Butane	Hexane	0.1-0.9	3.895-3.181	[21]
26	Butane	Heptane	0.1-0.9401	4.11-2.97	[21]
27	Butane	Octane	0.1-0.9	4.322-2.78	[21]
28	Butane	Decane	0.739-0.956	4.92-4.25	[21]
29	Pentane	Hexane	0.1685-0.8632	3.35-3.1	[32]
30	Pentane	Heptane	0.255-0.747	3.32-3.06	[21]
31	Pentane	Nonane	0.1-0.9	3.657-2.501	[21]
32	Hexane	Heptane	0.1-0.9	3.03-2.796	[21]
33	Hexane	Octane	0.1-0.9	3.052-2.595	[21]
34	Hexane	Decane	0.1-0.9	3.19-2.306	[21]
35	Hexane	Tridecane	0.1-0.9	3.516-1.99	[21]
36	Hexane	Tetradecane	0.1-0.9	3.668-1.757	[21]
37	Heptane	Octane	0.1-0.9	2.744-2.546	[21]
38	Nonane	Tridecane	0.1-0.9	2.322-1.822	[21]
39	Decane	Dodecane	0.1-0.9	2.126-1.904	[21]
40	Methane	Cyclohexane	0.608-0.765	28.19-20.06	[21]
41	Ethane	Cyclohexane	0.1378-0.8617	9.21-5.22	[21]
42	Pentane	Cyclopentane	0.1-0.9	4.391-3.493	[21]
43	Hexane	Cyclopentane	0.1-0.9	4.328-3.159	[21]
44	Hexane	Cyclohexane	0.1-0.9	3.961-3.128	[21]
45	Heptane	Cyclopentane	0.1-0.9	4.332-2.914	[21]
46	Heptane	Cyclohexane	0-1	3.97-2.761	[21,25]
47	Octane	Cyclopentane	0.1-0.9	4.375-2.7	[21]
48	Octane	Cyclohexane	0.1-0.9	3.903-2.639	[21]
49	Nonane	Cyclopentane	0.1-0.9	4.462-2.515	[21]
50	Nonane	Cyclohexane	0.1-0.9	3.919-2.444	[21]
51	Decane	Cyclohexane	0.1-0.9	3.977-2.314	[21]
52	Tridecane	Cyclohexane	0.1-0.9	4.213-1.978	[21]
53	Methane	2-methylpropane	0.337-0.69	11.57-7.1	[21]
54	Methane	2-methylbutane	0.15-0.3339	7.573-5.002	[21]
55	Methane	Neopentane	0-.0.845	25.5-17.83	[21]
56	Methane	Methylcyclohexane	0.91-0.93	2.679-2.566	[21]
57	Methane	Ethylene	0.11-0.85	6.74-5.17	[21]
58	Ethane	Ethylene	0.285-0.44	5.2-5.19	[21]
59	Ethane	Propene	0.0525-0.9502	5.118-4.662	[21]
60	Ethane	Acetylene	0.27-0.68	5.698-5.275	[21]
61	Ethane	Benzene	0.0646-0.8756	9.522-5.99	[21]
62	Propane	2-methylbutane	0.101-0.899	4.57-3.643	[21]
63	Propane	Acetylene	0.166-0.86	6.42-5.38	[21]
64	Propane	Benzene	0-1	6.74-4.41	[21,27]
65	Propane	2,2-dimethylbutane	0.1527-0.9194	4.698-3.517	[21]
66	Propane	2,3-dimethylbutane	0.1516-0.9153	4.836-3.615	[21]
67	Propane	2-methylpentane	0.1497-0.919	4.8-3.477	[21]

Table 3
Critical pressure, mole fraction, and reference of selected binary mixtures.

NO.	Component (1)	Component (2)	Range of x_1	Range of P_C (MPa)	Ref.
68	Propane	3-methylpentane	0.1451-0.885	4.901-3.613	[21]
69	Propane	Cumene	0-1	8.04-3.3	[27]
70	Pentane	Ethylbenzene	0.1-0.9	4.09-3.635	[21]
71	Hexane	Methylcyclopentane	0.1-0.9	3.707-3.17	[21]
72	Hexane	Cis-Decalin	0.1-0.9	4.027-3.436	[21]
73	Hexane	Acetylene	0.335-0.625	10.4-6.23	[21]
74	Hexane	Benzene	0.1-0.9	4.66-3.172	[21]
75	Hexane	Toluene	0.366-0.7	3.951-3.537	[21]
76	Hexane	Ethylbenzene	0.1-0.9	3.652-3.166	[21]
77	Heptane	Ethylene	0.0301-0.9296	10.45-3.25	[21]
78	Heptane	Benzene	0.1-0.9	4.584-2.894	[21]
79	Heptane	Ethylbenzene	0.1-0.9	3.561-2.845	[21]
80	Octane	Benzene	0.1-0.9	4.551-2.679	[21]
81	Octane	Ethylbenzene	0.1-0.9	3.485-2.593	[21]
82	Nonane	Benzene	0.1-0.9	4.553-2.494	[21]
83	Decane	Benzene	0.1-0.9	4.597-2.357	[21]
84	Tridecane	Benzene	0.1-0.9	4.853-1.986	[21]
85	Hexadecane	Benzene	0.1-0.9	5.257-1.792	[21]
86	Cyclopentane	Cyclohexane	0.1-0.9	4.505-4.154	[21]
87	Cyclopentane	Methylcyclopentane	0.1-0.9	4.459-3.872	[21]
88	Cyclopentane	Methylcyclohexane	0.1-0.9	4.49-3.64	[21]
89	Cyclohexane	Methylcyclopentane	0.1-0.9	4.062-3.828	[21]
90	Cyclohexane	Methylcyclohexane	0.1-0.9	4.035-3.559	[21]
91	Cyclohexane	Cis-Decalin	0.1-0.9	4.328-3.466	[21]
92	Methylcyclopentane	Methylcyclohexane	0.1-0.9	3.791-3.545	[21]
93	Cis-Decalin	Benzene	0.1-0.9	5.063-3.547	[21]
94	Cis-Decalin	Ethylbenzene	0.1-0.9	3.733-3.375	[21]
95	Cis-Decalin	O-Xylene	0.1-0.9	3.782-3.346	[21]
96	2,2-Dimethylbutane	2,2-Dimethylpentane	0.1-0.9	3.079-2.84	[21]
97	2,2-Dimethylbutane	2,2-Dimethylhexane	0.1-0.9	3.106-2.644	[21]
98	2,2-Dimethylbutane	2,2-Dimethylheptane	0.1-0.9	3.164-2.497	[21]
99	2,2-Dimethylpentane	2,2-Dimethylhexane	0.1-0.9	2.792-2.588	[21]
100	2,2-Dimethylpentane	2,2-Dimethylheptane	0.1-0.9	2.795-2.413	[21]
101	2,2-Dimethylhexane	2,2-Dimethylheptane	0.1-0.9	2.551-2.347	[21]
102	2,2-Dimethylpentane	2-Methylhexane	0.1-0.9	3.038-2.79	[21]
103	2,2-Dimethylpentane	2-Methylheptane	0.1-0.9	3.073-2.601	[21]
104	2,2-Dimethylpentane	2-Methyloctane	0.1-0.9	3.091-2.435	[21]
105	2-Methylhexane	2-Methylheptane	0.1-0.9	2.746-2.539	[21]
106	2-Methylhexane	2-Methyloctane	0.1-0.9	2.749-2.381	[21]
107	2-Methylheptane	2-Methyloctane	0.1-0.9	2.51-2.343	[21]
108	Ethylene	Propene	0.117-0.914	5.47-4.83	[21]
109	Ethylene	Acetylene	0.217-0.82	5.67-5.23	[21]
110	Propene	Butene	0.114-0.885	4.7-4.35	[21]
111	Propene	Acetylene	0.1-0.9	6.27-5.44	[21]
112	Benzene	Cyclopentane	0.1-0.9	4.903-4.566	[21]
113	Benzene	Cyclohexane	0.1-0.9	4.797-4.145	[21]
114	Benzene	Methylcyclopentane	0.1-0.9	4.765-3.879	[21]
115	Benzene	Methylcyclohexane	0.1-0.9	4.728-9.916	[21]
116	Benzene	Ethylene	0-1	11.53-5.01	[21,28]
117	Benzene	Toluene	0.1-0.9	4.836-4.2	[21]
118	Benzene	Ethylbenzene	0.1-0.9	4.823-3.767	[21]
119	Cumene	Benzene	0-1	4.89-3.3	[27]
120	Benzene	O-Xylene	0.1-0.9	4.865-3.881	[21]
121	Toluene	Ethylbenzene	0.1-0.9	4.07-3.666	[21]
122	Toluene	O-Xylene	0.1-0.9	4.102-3.772	[21]
123	O-Xylene	Ethylbenzene	0.1-0.9	3.706-3.626	[21]

3.2. ANN model

As discussed in Section 2.1, artificial neural networks are able to produce flexible models for complex systems with interconnecting inputs and outputs. Thus, after collecting the database, the next step is creating the neural network structure. In this study, the multilayer perceptron network (MLP) was used due to its flexibility and accuracy, which were reported by several researchers. The neural network designed to find the best network architecture and parameters is firstly taught by this database. Then, the trained network is tested using the test data intended to determine the validity of the data obtained from the network. In this work, 70 % of the data is used for training, 30 % for testing, and 10 % for validation. To increase the convergence rate and reduce the prediction error, all of the input data are normalized in the range [0.1-0.9] using the following formula:

$$(\text{Scaled Value}) = \frac{(\text{Actual Value}) - (\text{Minimum actual value})}{(\text{Maximum actual value}) - (\text{Minimum actual value})} \times 0.8 + 0.1 \quad (1)$$

The designed network is a feed-forward network with an error-back propagation algorithm based on Levenberg-Marquardt backward training release function. In this network, the sigmoid activation function is used for hidden and output layers. The designed neural network block diagram is shown in Fig. 1. In this network, inputs are given to the input layer and, then, are delivered to the hidden layer. Mathematical calculations are performed in the hidden layer using a particular transfer function and, eventually, are sent to the output layer and displayed as outputs.

The general structure of a neuron is shown in Fig. 2. Except for the neurons in the input layer, the other neurons receive a set of inputs and generate the desired output using the transfer function.

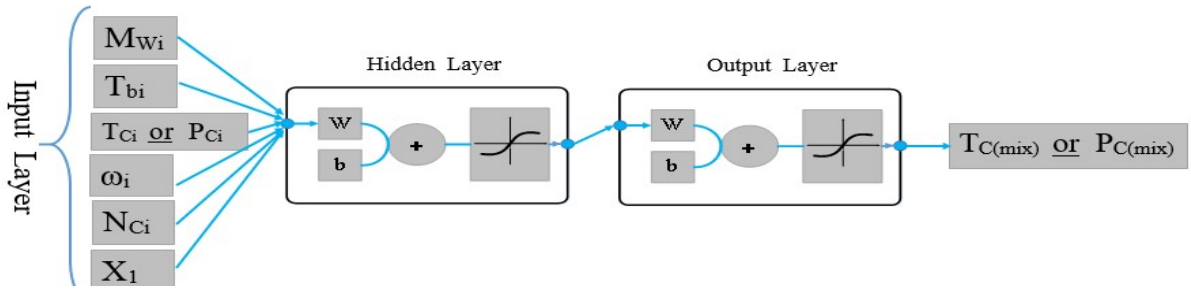


Figure 1. The block diagram of the used neural network.

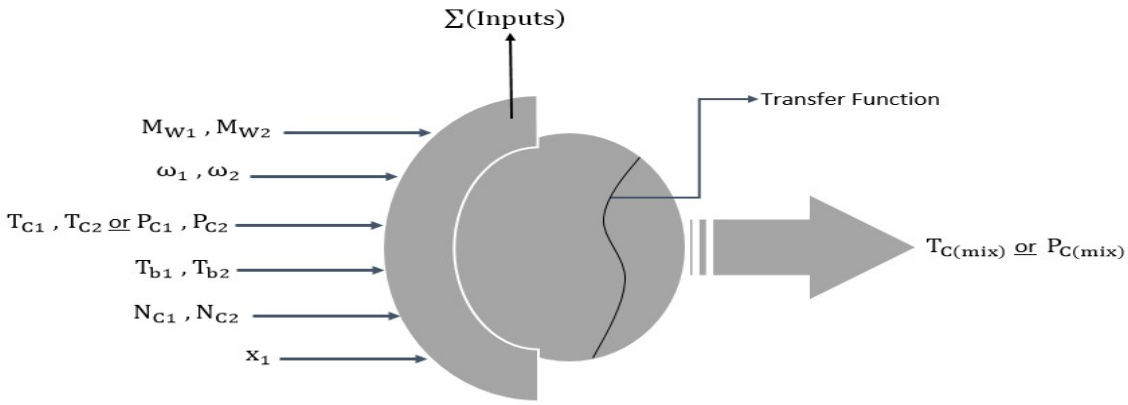


Figure 2. General structure of a neuron.

It is necessary to determine the optimal neuron number in the hidden layer to reach the optimal ANN model. Since the best performance of the artificial neural networks occurs in a particular and short range of neurons, a comprehensive search method can be used to determine the optimal network. Therefore, after introducing input and output data into the network and normalizing the data, we have considered 25 neurons as the maximum number of neurons. The application has been executed for 1 to 25 neurons, and the results are stored. The network training for each neuron is repeated for 20 times, and the average output is considered as the final output of the network. At the end of the application, the output data are saved, and the average relative deviation (ARD) is plotted in terms of the number of neurons in the hidden layer. The minimum ARD value of this chart determines the optimum neuron value. Performance of the model is measured by measuring the error of training, accuracy, and testing. The results are reported in the next section. The flowchart of ANN in this work is shown in Fig. 3.

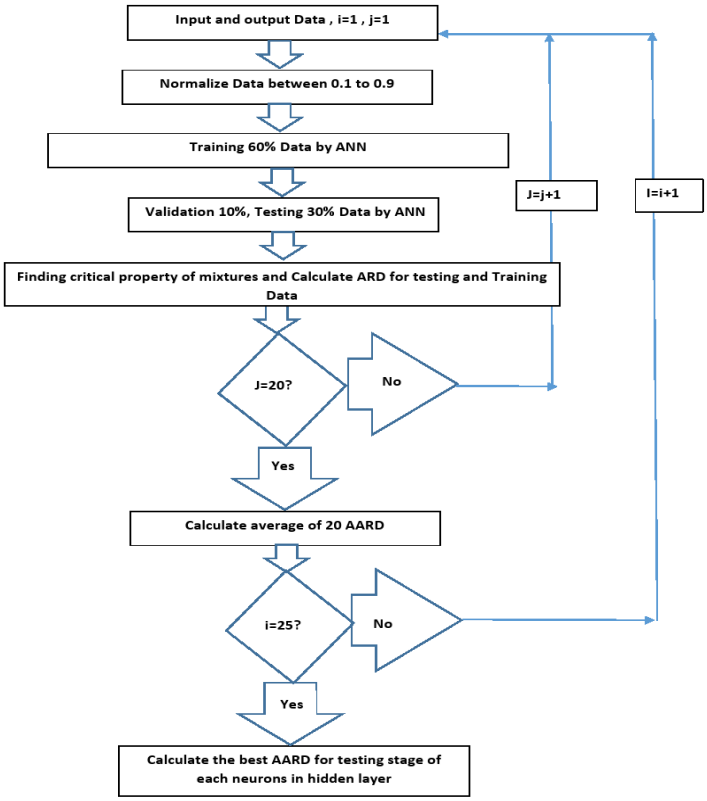


Figure 3. The flowchart of ANN for predicting critical properties of binary mixtures.

3.3. SVM-PSO model

The support vector regression introduced in Section 2-2 is selected as an alternative model for the prediction of true critical temperature and pressure of binary hydrocarbon mixtures.

The same data used in the ANN model are used to train and test the SVM model. It should be noted that the division of the

database into training and testing data has been done without overlapping. The inputs are selected and have been given to the network after normalization. The open-source LS-SVMlab toolbox that is introduced elsewhere [33] has been used for model generation. The generated SVM model is based on the radial basis function (RFB). Only two parameters of epsilon (the eigenvalue of compilation of complexity and model accuracy) and sigma (SD or width of the RBF) have to be determined in this kind of network [33]. To optimize these two parameters, the PSO algorithm is used and briefly described in Section 2-3.

The existing model in the training part is used to establish the ultimate model for the test data in SVM. The Kernels are then used in data separation, indicating non-linear and suitable mapping models. The training data are used to predict the real values in the output of the network. All of the regression problems cannot be regressed linearly; therefore, kernels may be used properly. Radial basis function and Polynomial kernel are considered as two of the most widely used kernels. Selection of the appropriate kernels and fitting the parameters can transform a non-linear separable problem into a linearly separable one. In this work, we have used the radial basis function as a proper kernel for SVM. PSO is an optimization procedure used successfully for reasons such as implementation, perception, and simple correction. GA is another technique that is used to generate a new generation from the previous one in the PSO. In PSO, the position and velocity start if each particle changes randomly. The position in each iteration depends on the velocity and previous location [34].

Data was randomly divided into 60 % for training and 30 % for test and 10 % for validation in the SVM prediction. All the data were placed between -1 and 1 by the following formula:

$$\text{xnew} = 2 * \frac{\text{xold}-\text{xmin}}{\text{xmax}-\text{xmin}} - 1 \qquad (2)$$

Three parameters of c, σ, and ε were obtained in SVM by PSO optimization for training data; then, these parameters were used to predict critical pressure and temperature for the test data. The procedure for SVM-PSO algorithm is showed in Figure 4.

3.4. Performance criteria for two models

The accuracy and precision of each model are examined with different measurement criteria. Different statistical criteria including the average relative deviation (ARD), relative absolute error (RAE), coefficient of determination (R^2), mean absolute error (MAE), and mean squared error (MSE) are defined according to the following equations.

$$\text{MAE} = \frac{1}{\text{N}}\sum_{\text{i=1}}^{\text{N}} \left|\text{y}_{\text{C}_\text{i}}^{\text{pred}} - \text{y}_{\text{C}_\text{i}}^{\text{exp}}\right| \qquad (3)$$

$$\text{ARD}(\%) = \frac{1}{\text{N}} \sum_{\text{i=1}}^{\text{N}} 100 \times \left|\frac{\text{y}_{\text{C}_\text{i}}^{\text{exp}} - \text{y}_{\text{C}_\text{i}}^{\text{pred}}}{\text{y}_{\text{C}_\text{i}}^{\text{exp}}}\right| \qquad (4)$$

$$\text{RAE}(\%) = \frac{\sum_{\text{i=1}}^{\text{N}}\left|\text{y}_{\text{C}_\text{i}}^{\text{pred}} - \text{y}_{\text{C}_\text{i}}^{\text{exp}}\right|}{\sum_{\text{i=1}}^{\text{N}}\left|\text{y}_{\text{C}_\text{i}}^{\text{exp}} - \bar{\text{y}}^{\text{exp}}\right|}, \bar{\text{y}}^{\text{exp}} = \frac{1}{\text{N}}\sum_{\text{i=1}}^{\text{N}} \text{y}_{\text{C}_\text{i}}^{\text{exp}} \qquad (5)$$

$$\text{MSE} = \frac{1}{\text{N}}\sum_{\text{i=1}}^{\text{N}} (\text{y}_{\text{C}_\text{i}}^{\text{pred}} - \text{y}_{\text{C}_\text{i}}^{\text{exp}})^2 \qquad (6)$$

$$R^2 = \left(\frac{N\sum_{i=1}^{N}\left(\text{y}_{\text{C}_\text{i}}^{\text{pred}}.\text{y}_{\text{C}_\text{i}}^{\text{exp}}\right) - \left(\sum_{i=1}^{N}\text{y}_{\text{C}_\text{i}}^{\text{exp}}\right)\left(\sum_{i=1}^{N}\text{y}_{\text{C}_\text{i}}^{\text{pred}}\right)}{\left[N\sum_{i=1}^{N}\left(\text{y}_{\text{C}_\text{i}}^{\text{exp}}\right)^2 - \left(\sum_{N}^{i=1}\text{y}_{\text{C}_\text{i}}^{\text{exp}}\right)^2\right]\left[N\sum_{i=1}^{N}\left(\text{y}_{\text{C}_\text{i}}^{\text{pred}}\right)^2 - \left(\sum_{N}^{i=1}\text{y}_{\text{C}_\text{i}}^{\text{pred}}\right)^2\right]}\right)^2 \qquad (7)$$

where $\text{y}_{\text{c}_\text{i}}^{\text{exp}}$ is the experimental output for the i^{th} mixture, $\text{y}_{\text{c}_\text{i}}^{\text{pred}}$ is the predicted output by ANN or the SVM, and N is the number of experimental points available for the critical temperature or pressure.

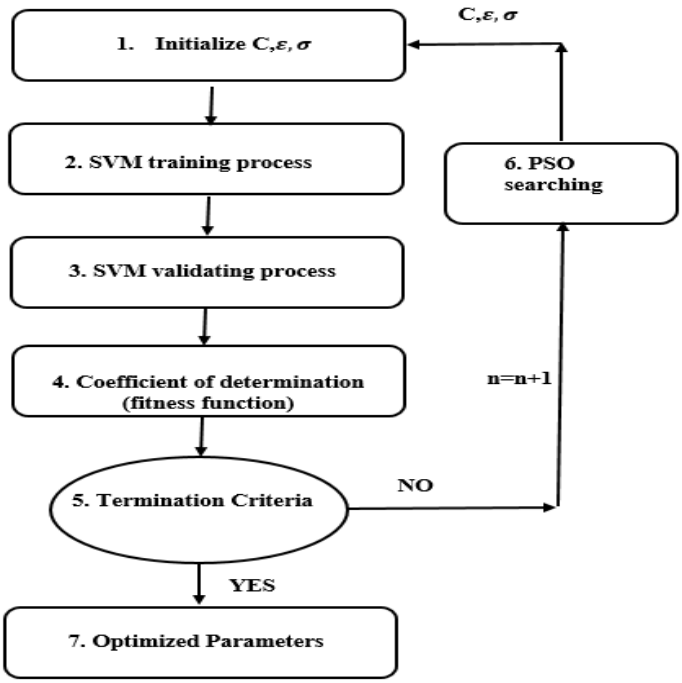


Figure 4. Flowchart of the PSO–SVM-based model (Optimization of SVM parameters and critical pressure and temperature estimation)[34].

4. Results and discussion

4.1. Results of artificial neural networks algorithm

Two datasets have been collected for critical temperature and pressure of binary hydrocarbon mixtures, respectively. The first dataset contains 163 binary mixtures with 1646 data points, and the second database contains 124 binary mixtures with 1090 data points. As stated in Section 3.2, 60 % of the data are used for training, 30 % for network testing, and 10 % for network validation. After training the network and creating the outputs, the average error is calculated using certain statistical criteria.

The designed network is a feed-forward network with an error-back propagation algorithm, the Levenberg-Marquardt backward training release function, and the sigmoid activation function for the hidden layer and the output layer. The ARD between the model outputs and the testing experimental data is calculated for different number of neurons in the hidden layer and is reported in Table 4 and plotted in Fig. 5.

Table 4
The results of average relative deviation (ARD) in different neurons.

Hidden neuron	Error analysis	
	$T_{C(mix)}$	$P_{C(mix)}$
	ARD(%)	ARD(%)
1	0.0991	0.1351
2	0.2112	0.6215
3	0.0267	0.0874
4	0.0209	0.0749
5	0.0194	0.3182
6	0.0184	0.0483
7	0.0732	0.0451
8	0.0172	0.0459
9	**0.0161**	0.0406
10	0.0694	0.0397
11	0.0179	**0.0387**
12	0.0693	0.0489
13	0.0220	0.0436
14	0.0763	0.0490
15	0.0196	0.3096
16	0.0204	0.0440
17	0.0217	0.0649
18	0.0196	0.0604
19	0.0212	0.3173
20	0.0292	0.0412
21	0.0227	0.0577
22	0.0242	0.0559
23	0.0253	0.0504
24	0.0230	0.0706
25	0.0262	0.0733

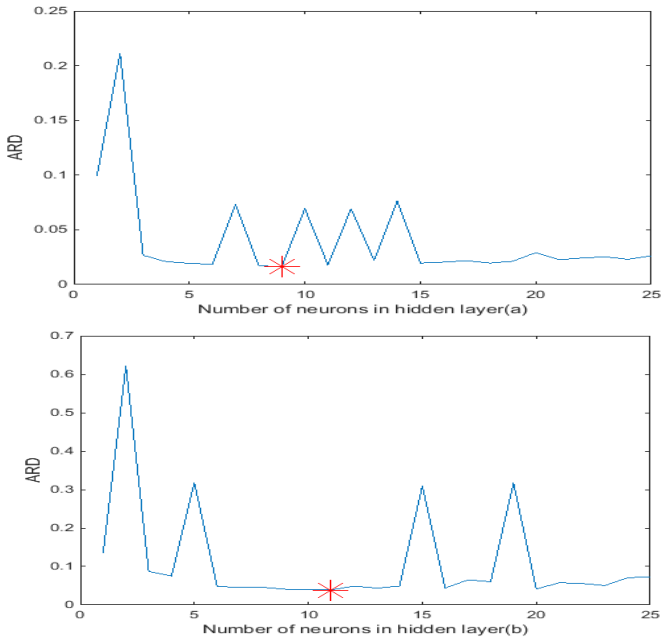


Figure 5. Deviation of ARD with different number of neurons in the hidden layer: (a) the true critical temperature, (b) the true critical pressure.

It can be easily observed that the optimum ANN architecture contains 9 and 11 in the hidden layer for the prediction of critical temperature and pressure, respectively. The final ANN models designed to predict the true critical temperature and pressure of the binary hydrocarbon mixtures are shown in Fig. 6.

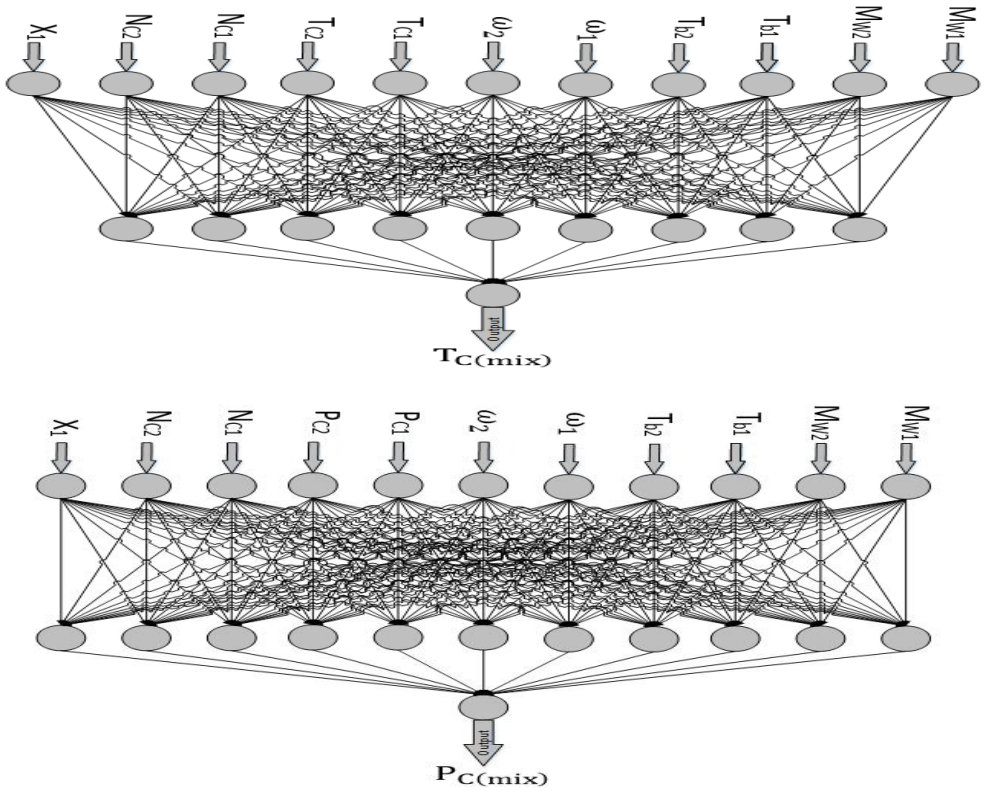


Figure 6. General structure of the neural network used in this work.

Squared error (MSE) in optimal neuron is shown in Table 5. Figures 7 and 8 show the training and testing results of the designed neural network for prediction of the true critical temperature and pressure of the binary hydrocarbon mixtures.

Table 5.
Statistical criteria for the prediction of the true critical temperature and pressure of binary hydrocarbon mixtures by artificial neural networks.

	$T_{C(mix)}$					$P_{C(mix)}$				
	ARD	**MAE**	**RAE**	**MSE**	**R^2**	**ARD**	**MAE**	**RAE**	**MSE**	**R^2**
Training	0.0103	0.0043	0.0324	6.7053×10^{-5}	0.9975	0.0130	0.0025	0.0441	2.2877×10^{-5}	0.9981
Testing	0.0161	0.0064	0.0475	2.5165×10^{-4}	0.9907	0.0387	0.0087	0.1505	0.0020	0.9635
Total	0.0134	0.0051	0.0394	4.6112×10^{-4}	0.9938	0.0195	0.0062	0.0941	1.1822×10^{-5}	0.9978

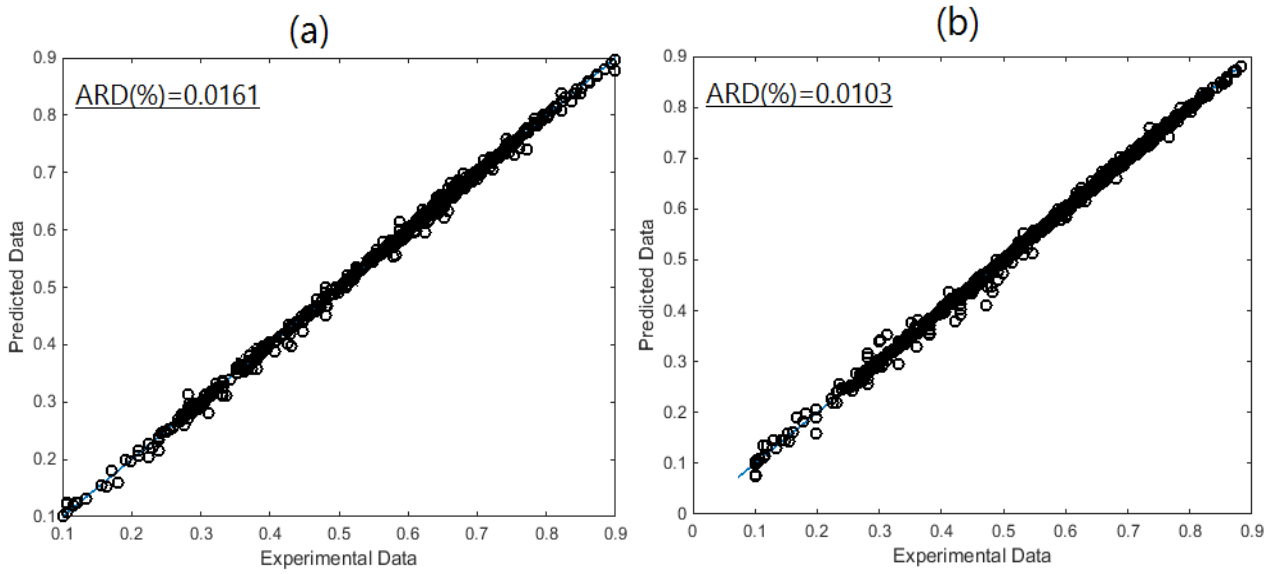


Figure 7. Comparison of experimental and predicted data for the prediction of true critical temperature of binary hydrocarbon mixtures by ANN based model: (a) for testing data and (b) for training data.

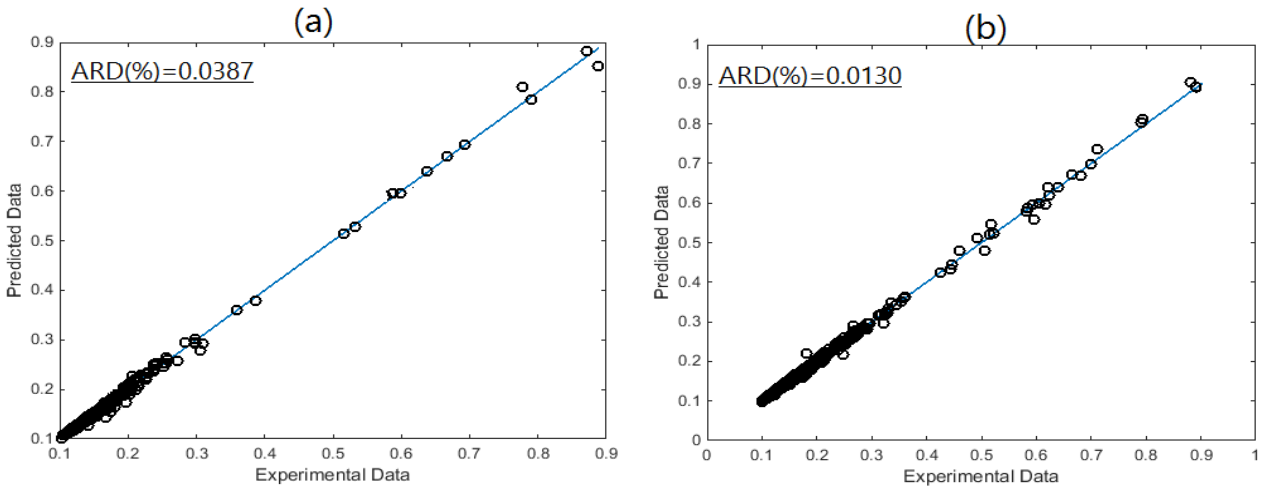
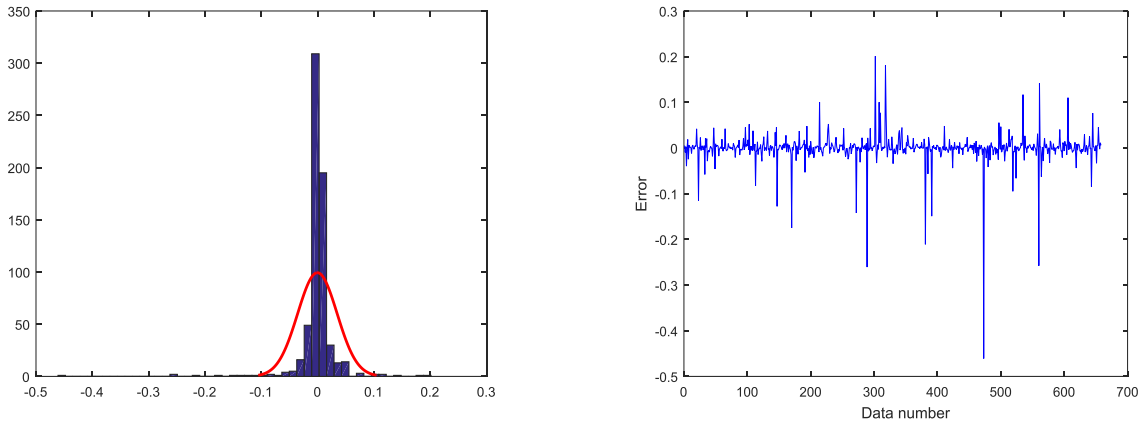


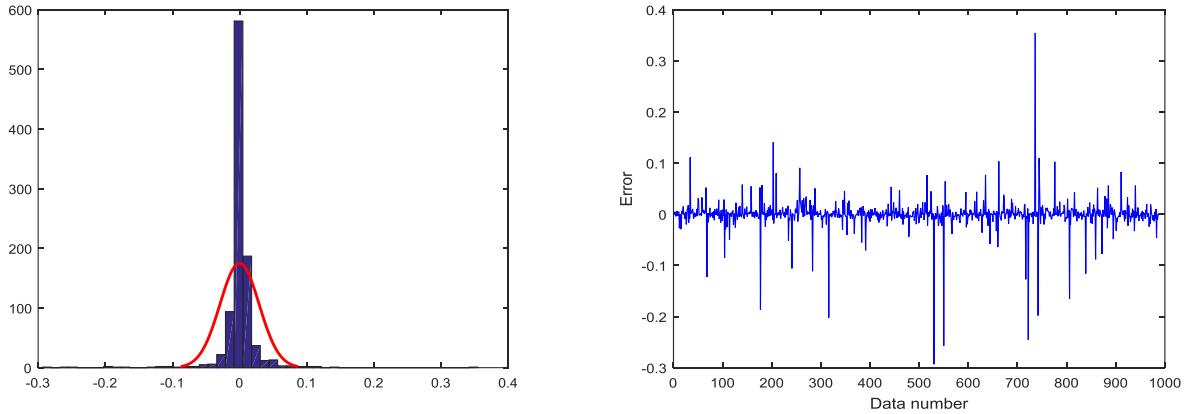
Figure 8. Comparison of experimental and predicted data for the prediction of true critical pressure of binary hydrocarbon mixtures of ANN-based model: (a) the testing result and (b) the training result.

Section A is related to the prediction of the true critical temperature of the binary hydrocarbon mixtures, and part B is related to the true critical pressure of the binary hydrocarbon mixtures. It should be noted that the designed network is tested using some experimental data, which were not used in the training phase. The circular points show the critical temperature and pressure predicted by the network, and the 45-degree line in the middle of the diagrams represents the exact identity of the experimental and predicted values. In other words, the more circular points are located near the 45-degree line, and the accuracy of the network is higher. Therefore, it seems that the designed ANN model is able to predict the critical temperature and pressure of binary hydrocarbon mixtures. The ARD for critical temperature in the training and test phases is 0.0103 and 0.0161, respectively. These values are equal to 0.0130 and 0.0387 for critical pressure, respectively. These numbers indicate the acceptable accuracy of the ANN model for this case. Histogram plot and data sensitivity plot for testing, validation, and training data ARD of critical temperature and critical pressure prediction by ANN are shown in Fig. 9 to Fig. 12.



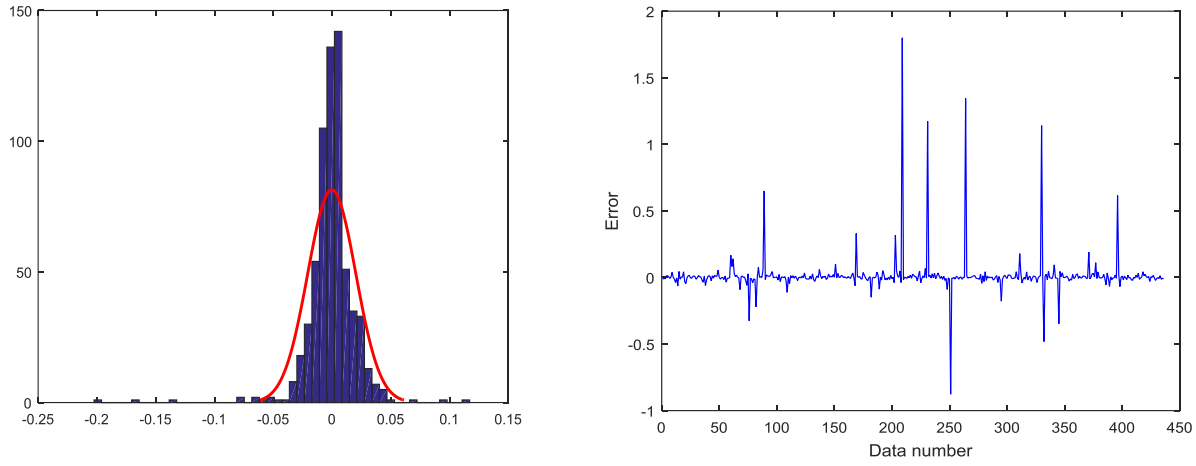
(a) (b)

Figure 9. Histogram plot (a) and data sensitivity analysis graph (b) for test and validation data ARD of critical temperature prediction by ANN.



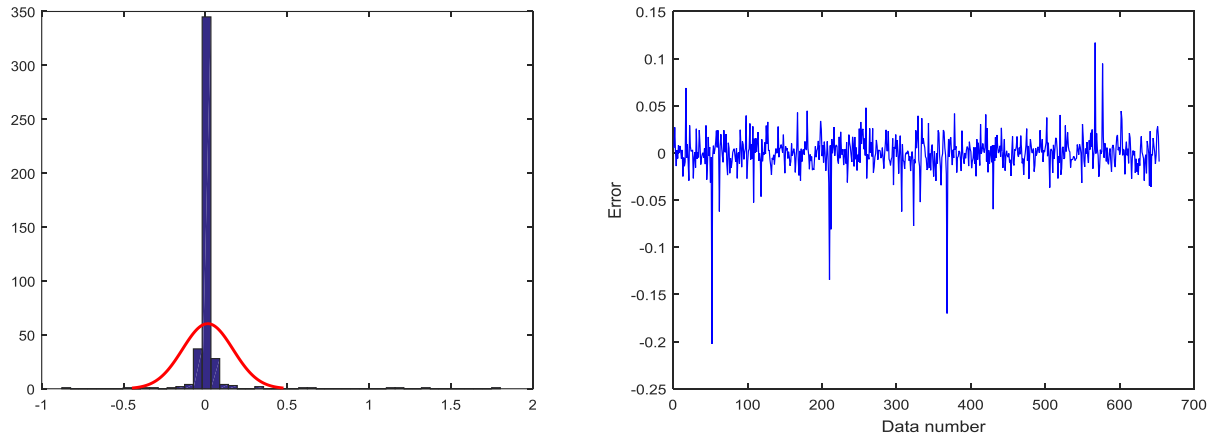
(a) (b)

Figure 10. Histogram plot (a) and data sensitivity analysis graph (b) for train data ARD of critical temperature prediction by ANN.



(a) (b)

Figure 11. Histogram plot (a) and data sensibility analysis graph (b) for test and validation data ARD of critical pressure prediction by ANN.



(a) (b)

Figure 12. Histogram plot (a) and data sensitivity analysis graph (b) for train data ARD of critical pressure prediction by ANN.

As shown in Figs. 9 to 12, Gaussian plot of errors is shown with the red curve in histogram of ARD errors. Variance and mean of ARD for testing and training of critical temperature are (0.0005, -0.00039) and (0.00025,-0.0002) and for testing and training of critical pressure are (0.00097, 0.002) and (0.000078, -0.000124), respectively. In all of these parts, means of ARD are near zero. However, variance of critical pressure in test part is larger than that of the others. It is shown that ARD variation range in the test part of critical pressure calculation is more than the others. Less variance of ARD shows more accurate prediction of ANN in each part.

4.2. Results of SVM-PSO algorithm

Modeling with support vector regression is conducted using the same inputs of the neural network. After normalizations of the experimental data, 70 % of the data are provided for training and 30 % for testing. Since the radial basis kernel has a good ability to predict, this function is used in this model. Before model training in SVM, two main parameters of sigma and epsilon have to be optimized to reach the maximum accuracy.

The particle swarm optimization algorithm (PSO) is used in order to optimize the two parameters of sigma and epsilon. Finally, the statistical error parameters are calculated to analyze the model's accuracy. Table 6 shows the values of two parameters obtained by PSO to predict the true critical temperature and pressure of the binary hydrocarbon mixtures. The prediction results with support vector regression are presented in Figures 13 and 14.

Table 6
The optimized values of SVM model obtained by PSO algorithm.

Parameter	$T_{C(mix)}$	$P_{C(mix)}$
Kernel function	RBF	RBF
Epsilon	3.7386	1.3065
Sigma	5.6994×10^{3}	6.581×10^{3}

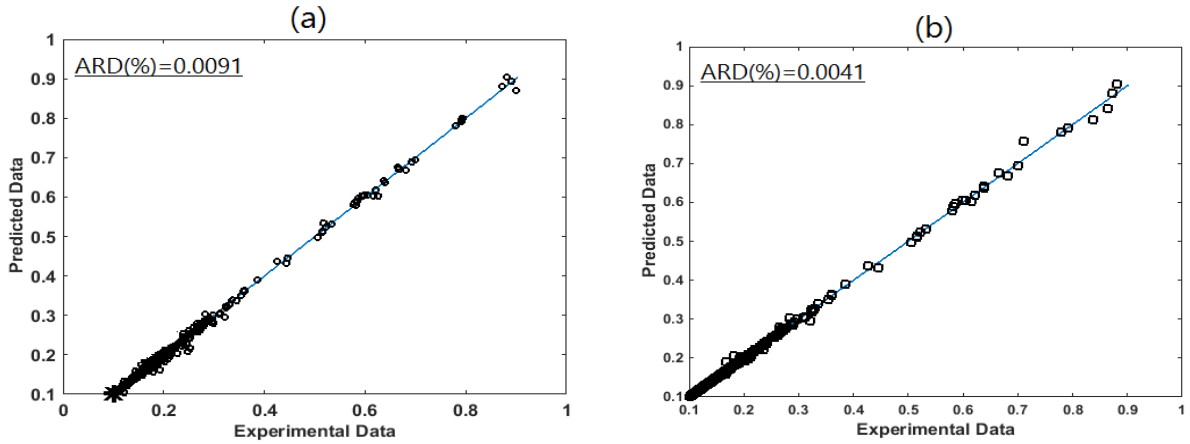


Figure 13. Comparison of experimental and predicted data to predict the true critical temperature of binary hydrocarbon mixtures by SVR model: (a) the testing results and (b) the training results.

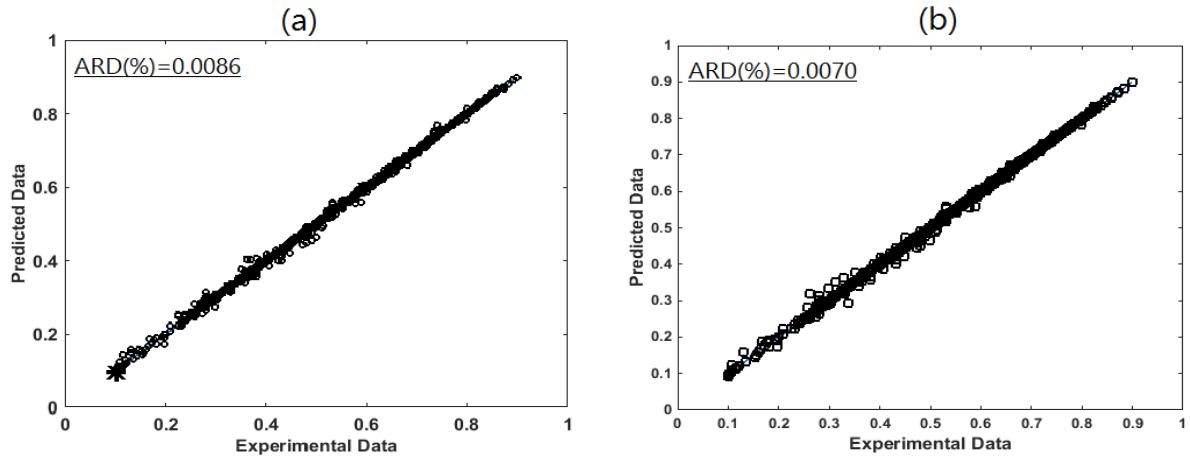


Figure 14. Comparison of experimental and predicted data for prediction of true critical pressure of binary hydrocarbon mixtures by SVR model: (a) the testing results and (b) the training results.

Section A of these figures is related to the training phase, and section B is related to the testing phase. The diagonal line represents the equality of the predited results and the experimental data. It can be easily observed that the model predictions are located near the diagonal line, meaning that the values predicted by the SVM are very close to the experimental data, indicating the high accuracy of the support vector regression model. The matching of experimental and predicted data in this model shows that the

SVM model with high accuracy can predict the true critical temperature and pressure of binary hydrocarbon mixtures. The average relative deviation (ARD), relative absolute error (RAE), coefficient of determination (R^2), mean absolute error (MAE), and mean squared error (MSE) constitute the statistical criteria used to show the performance and efficiency of the SVM model and are presented in Table 7.

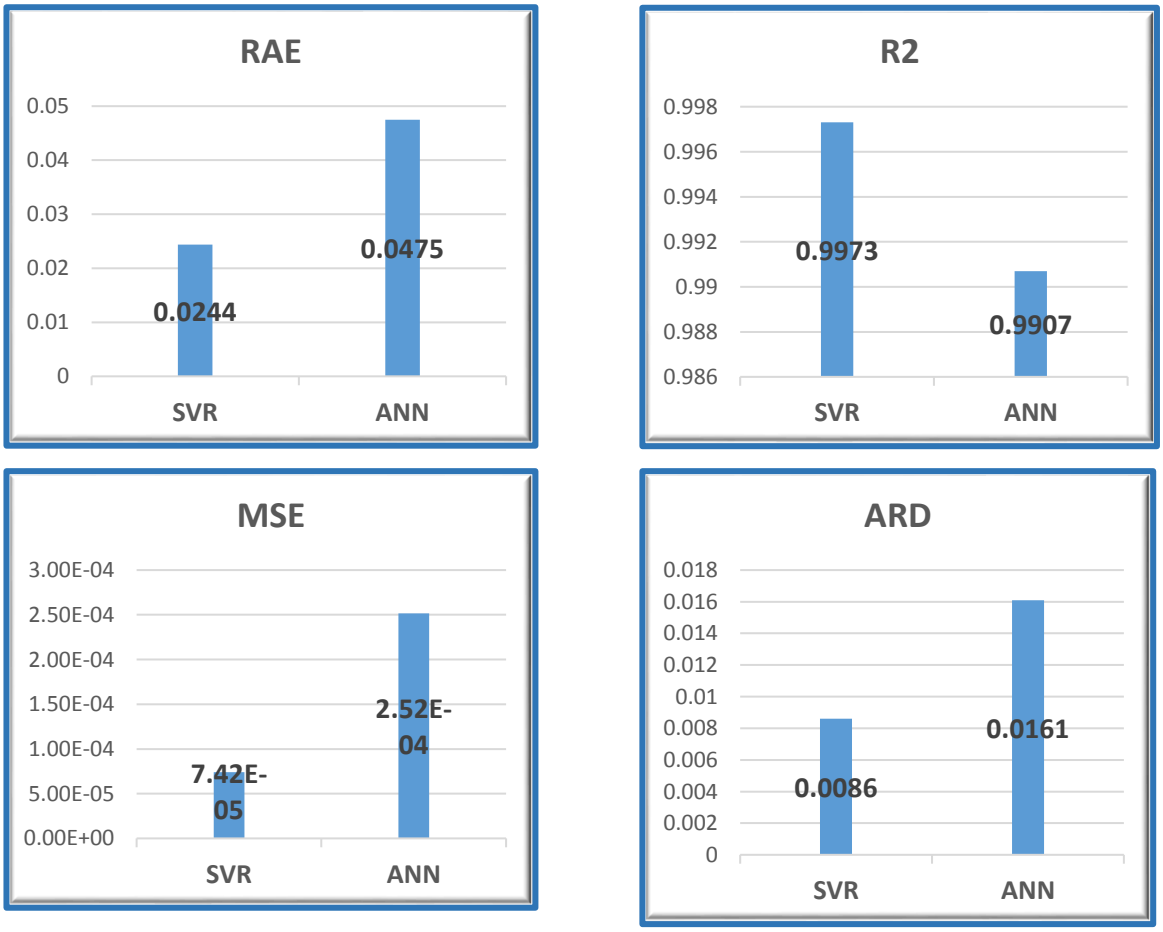
Table 7
Statistical measures for prediction of the true critical temperature and pressure of binary hydrocarbon mixtures by the SVM model.

	$T_{C(mix)}$					$P_{C(mix)}$				
	ARD	**MAE**	**RAE**	**MSE**	**R^2**	**ARD**	**MAE**	**RAE**	**MSE**	**R^2**
Training	0.0070	0.0026	0.0221	4.3290×10^{-5}	0.9984	0.0041	0.0012	0.0205	2.0388×10^{-5}	0.9984
Testing	0.0086	0.0030	0.0244	7.4161×10^{-5}	0.9973	0.0091	0.0022	0.0407	5.8069×10^{-5}	0.9953
Total	0.0110	0.0043	0.0391	1.2040×10^{-4}	0.9956	0.0187	00041	0.0722	1.1450×10^{-5}	0.9981

According to the results of Tables 5 and 7, it is clearly seen that the values of the statistical parameters for the SVM model are less than those of the ANN model, indicating the higher accuracy and reliability of the SVM method. In addition, support vector regression is preferable to neural networks due to the high speed and precision.

4.3. Comparison of the SVM and ANN algorithms

The statistical error criteria are shown in Fig. 15 to compare the accuracy of the SVM-PSO with ANN models to predict the true critical temperature of binary hydrocarbon mixtures.



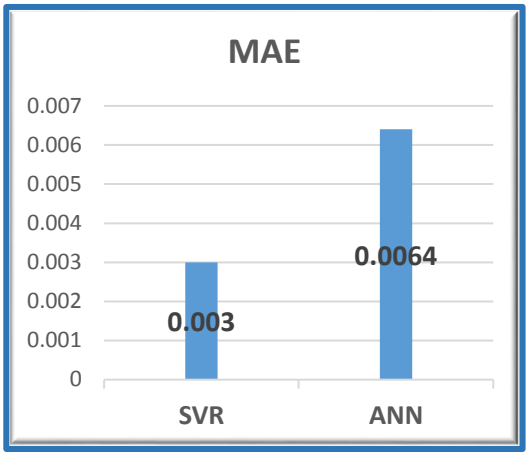
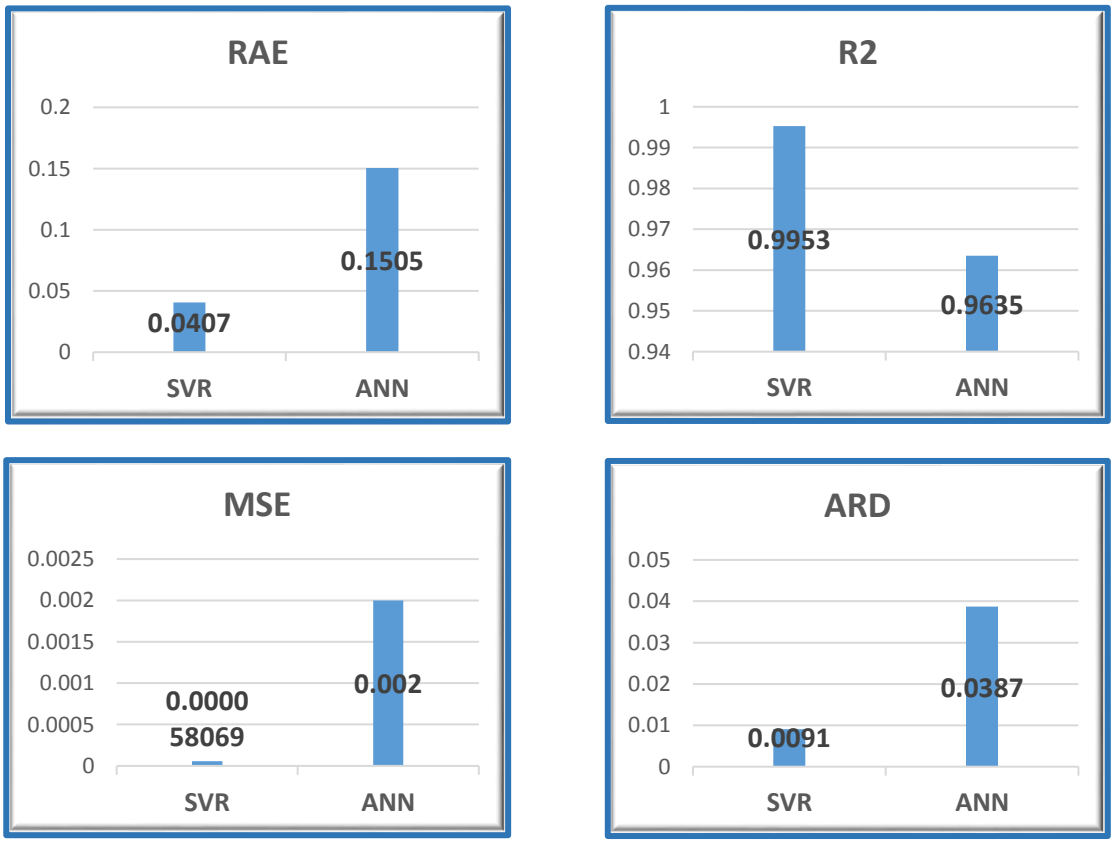


Figure 15. Statistical criteria for prediction of the true critical temperature of the binary hydrocarbon mixtures using ANN and SVR.

In this figure, the values of ARD, MAE, MSE, RAE, and R^2 parameters for the ANN model were equal to 0.161, 0.0064, 2.5165E-4, 0.0475, and 0.9907, respectively, and the values of ARD, MAE, MSE, RAE, and R^2 parameters for the SVM-PSO model were equal to 0.0086, 0.003, 7.4161E-5, 0.0244, and 0.9973, respectively. The obtained values confirm the high accuracy of both models to predict the critical temperature of the binary hydrocarbon mixtures. Moreover, the above values show that, in all of these parameters, the SVM-based model is always more accurate than the ANN-based model.

In Fig. 16, the statistical error criteria are shown to predict the true critical pressure of binary hydrocarbon mixtures for two models of ANN and SVM-PSO.



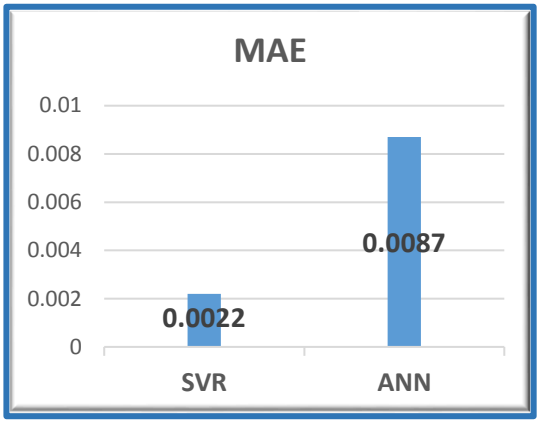


Figure 16. Statistical criteria for prediction of the true critical pressure of the binary hydrocarbon mixtures using ANN and SVR.

In this figure, the average relative deviation (ARD), relative absolute error (RAE), coefficient of determination (R^2), mean absolute error (MAE), and mean squared error (MSE) have been investigated. In the ANN model, the values of ARD, MAE, MSE, RAE, and R^2 parameters were 0.0387, 0.0087, 0.002, 0.1505, and 0.9635, respectively; 0in SVM model, the values of parameters were 0.0091, 0.0087, 5.8069E-5, 0.0407, and 0.9953, respectively. In this case, the obtained values also confirm the high accuracy of both models to predict the critical pressure of the binary hydrocarbon mixtures. In addition, the above values also show that the SVM-based model is always more accurate than the ANN-based model.

Therefore, it can be concluded that the SVM-PSO model is more accurate than the ANN model to predict the critical properties of binary hydrocarbon mixtures.

4.4. Comparison of the running speed of SVM-PSO and ANN algorithms

In Sections 4.1-4.3, the accuracy of each model was studied. However, in addition to accuracy, the runtime is also an important criterion for the selection of the model. The runtime values of these two models in a computer system with 4 Gigabytes of memory and an Intel core i5-2430 M-2.4 GHz processor are reported in Table 8. To provide a more precise timetable, each row in this table is executed for ten times, and its average time is reported in the table.

Table 8
Runtime for ANN and SVM models.

Model output	ANN runtime (s)	ANN runtime (s)
Critical temperature	18886	0.668
Critical pressure	15857	0.358

The table above shows that the support vector regression–based model supports learning and experimentation much faster than the artificial neural network model. Since, in accordance with the diagram in Section (4.3), the SVM-PSO provides higher precision. It can be concluded that SVM-PSO model will serve as the superior model for predicting critical temperature and pressure of binary mixtures.

5. Conclusions

Considering the importance of knowing the critical properties of multi-component mixtures without performing any laboratory processes was the motive of this study. As there are not many published works

concerning the application of soft computing models to predict the true critical properties of binary mixtures, two well-known models of MLP-ANN and SVM-PSO were selected in this study. The MLP-ANN model was optimized using a comprehensive search method and a SVM-based model using the particle swarm optimization algorithm. In addition, to carry out a realistic assessment, previous experimental data were used to train and measure the models. The statistical analysis results showed that although both models had high accuracy in predicting the critical temperature and pressure, the SVM-PSO model performed more accurate predictions than the MLP-ANN model. The SVM-PSO model training was much faster than the MLP-ANN model, making SVM-PSO the final choice for prediction of critical properties of binary hydrocarbon mixtures.

Acknowledgement
The authors acknowledge the funding support of Babol Noshirvani University of Technology through Grant Program No. BNUT/370675/98.

References

[1] He, M., Xin, N., Liu, Y. and Zhang, Y., "Determination of critical properties for binary and ternary mixtures of short chain alcohols and alkanes using a flow apparatus", *J. Supercrit. Fluids*, **104**, 19 (2015).

[2] Wang, L., Han, K., Xia, S., Ma, P. and Yan, F., "Measurement and correlation of critical properties for binary mixtures and ternary mixtures containing gasoline additives", *J. Chem. Thermodyn.*, **74**, 161 (2014).

[3] Belyakov, M. Y., Gorodetskii, E. E., Kulikov, V. D., Muratov, A. R., Voronov, V. P., Grigoriev, B. A. and Volkov, A. N., "Anomalous properties of dew-bubble curves in the vicinity of liquid-vapor critical points", *Fluid Phase Equilib.*, **358**, 91 (2013).

[4] Poling, B. E., Prausnitz, J. M. and O'Connell, J. P., The Properties of gases and liquids, fifth edition, New York, (2001).

[5] Najafi, H., Maghbooli, B. and Sobati, M. A., "Prediction of true critical temperature of multi-component mixtures: An extension to Chueh and Prausnitz method", *Fluid Phase Equilib.*, **363**, 1 (2014).

[6] Hosseini-Nasab, S. M., Manteghian, M., Sefti, M. V., Izadpanah, A. A. and Zare, M., "A neuro-fuzzy model as a predictive tool for the vapor-liquid equilibrium of binary mixtures", *Pet. Sci. Technol.*, **31**, 68 (2013).

[7] Ansari, H. R. and Gholami, A., "An improved support vector regression model for estimation of saturation pressure of crude oils", *Fluid Phase Equilib.*, **402**, 124 (2015).

[8] Najafi, H., Maghbooli, B. and Sobati, M. A., "Prediction of true critical temperature of multi-component mixtures: Extending fast estimation methods", *Fluid Phase Equilib.*, **392**, 104 (2015).

[9] Zhao, Y., Zhang, X., Deng, L. and Zhang, S., "Prediction of viscosity of imidazolium-based ionic liquids using MLR and SVM algorithms", *Comput. Chem. Eng.*, **92**, 37 (2016).

[10] Mehdizadeh, B. and Movagharnejad, K., "A comparison between neural network method and semi empirical equations to predict the solubility of different compounds in supercritical carbon dioxide", *Fluid Phase Equilib.*, **303**, 40

(2011).

[11] Hayer, H., Haghbakhsh, R., Keshtkari, S. and Raeissi, S., "Support vector machine and CPA EoS for the prediction of high-pressure liquid densities of normal alkanols", *J. Taiwan Inst. Chem. Eng.*, **45**, 2888 (2014).

[12] Lashkarbolooki, M., Hezave, A. Z. and Ayatollahi, S., "Artificial neural network as an applicable tool to predict the binary heat capacity of mixtures containing ionic liquids", *Fluid Phase Equilib.*, **324**, 102 (2012).

[13] Sabzevari, S. and Moosavi, M., "Density prediction of liquid alkali metals and their mixtures using an artificial neural network method over the whole liquid range", *Fluid Phase Equilib.*, **361**, 135 (2014).

[14] Fogel, D. B., Liu, D. and Keller, J. M., Fundamentals of computational intelligence, John Wiley & Sons, Inc., Hoboken, NJ, USA, (2016).

[15] Hemmati-Sarapardeh, A., Shokrollahi, A., Tatar, A., Gharagheizi, F., Mohammadi, A. H., Naseri, A., "Reservoir oil viscosity determination using a rigorous approach", *Fuel*, **116**, 39 (2014).

[16] dos Santos, L. C., Tavares, F. W., Ahón, V. R. R. and Kontogeorgis, G. M., "Modeling MEA with the CPA equation of state: A parameter estimation study adding local search to PSO algorithm", *Fluid Phase Equilib.*, **400**, 76 (2015).

[17] Panda, S. and Padhy, N. P., "Comparison of particle swarm optimization and genetic algorithm for FACTS-based controller design", *Appl. Soft Comput.*, **8**, 1418 (2008).

[18] Hassan, R., Cohanim, B., de Weck, O. and Venter, G., "A comparison of particle swarm optimization and the genetic algorithm", In: *46th AIAA/ASME/ASCE/AHS/ASC Struct. Struct. Dyn. Mater. Conf.*, American Institute of Aeronautics and Astronautics, Reston, Virigina, (2005).

[19] Poli, R., Kennedy, J. and Blackwell, T., "Particle swarm optimization", *Swarm Intell.*, **1**, 33 (2007).

[20] Kennedy, J. and Eberhart, R., "Particle swarm optimization", Proceeding of *ICNN'95 - Int. Conf. Neural Networks, IEEE*, n.d., pp. 1942-1948.

[21] API, Technical data book- Petroleum refining, (1997).

[22] Hicks, C. P. and Young, C. L., "Gas-liquid critical properties of binary mixtures", *Chem. Rev.*, **75**, 119 (1975).

[23] Belyakov, M. Y., Gorodetskii, E. E., Kulikov, V. D., Voronov, V. P. and Grigoriev, B.A., "Scaled equation of state and specific thermodynamic behavior of near-critical methane-pentane binary mixture", *Fluid Phase Equilib.*, **418**, 44 (2016).

[24] Nesterova, T. N., Vostrikov, S. V., Nesterov, I. A., Nazmutdinov, A. G. and Sosin, S. E., "Critical and maximum temperatures of coexistence of liquid and gas phase in hydrocarbons binary mixtures, I: Critical (vapour–liquid) temperatures of alkane binary mixtures", *Fluid Phase Equilib.*, **368**, 14 (2014).

[25] Teja, A. S. and Smith, R. L., "Critical properties of thermally unstable substances from mixture data", *AIChE J.*, **33**, 1560 (1987).

[26] Han, K., Xia, S., Ma, P., Yan, F. and Liu, T., "Measurement of critical temperatures and critical pressures for binary mixtures of methyl tert-butyl ether (MTBE)+alcohol and MTBE+alkane", *J.*

Chem. Thermodyn., **62**, 111 (2013).

[27] Vostrikov, S. V., Nesterova, T. N., Nesterov, I. A., Sosin, S. E. and Nazmutdinov, A. G., "III. Study of critical and maximum temperatures of coexistence of liquid and gas phase in hydrocarbons binary mixtures of aromatic hydrocarbons with alkanes and cycloalkanes", *Fluid Phase Equilib.*, **377**, 56 (2014).

[28] Wang, G., Qin, Z., Liu, J., Tian, Z., Hou, X. and Wang, J., "Critical properties of the reacting mixture in the alkylation of benzene with propene", *Ind. Eng. Chem. Res.*, **42**, 6531 (2003).

[29] Liu, T., Fu, J. -Y., Wang, K., Gao, Y. and Yuan, W. -K., "Gas-liquid critical properties of ethylene + benzene", *J. Chem. Eng. Data*, **46**, 809 (2001).

[30] Jones, I. W. and Rowlinson, J. S., "Gas-liquid critical temperatures of binary mixtures, Part 2", *Trans. Faraday Soc.*, **59**, 1702 (1963).

[31] Horstmann, S., Fischer, K. and Gmehling, J., "Experimental determination of critical data of mixtures and their relevance for the development of thermodynamic models", *Chem. Eng. Sci.*, **56**, 6905 (2001).

[32] Singh, H., Lucien, F. P. and Foster, N. R., "Critical properties for binary mixtures of ethane containing low concentrations of n –Alkane", *J. Chem. Eng. Data*, **45**, 131 (2000).

[33] Liu, J., Qin, Z., G. Wang, G., Hou, X. and Wang, J., "Critical properties of binary and ternary mixtures of hexane + methanol, hexane + carbon dioxide, methanol + carbon dioxide, and hexane + carbon dioxide + methanol", *J. Chem. Eng. Data*, **48**, 1610 (2003).

[34] Nieto, P. G., García-Gonzalo, E., Fernández, J. A. and Muñiz, C. D., "Hybrid PSO–SVM-based method for long-term forecasting of turbidity in the Nalón river basin: A case study in Northern Spain", *Ecological Engineering*, **73**, 192 (2014).